

## Worksheet Review : (p. 3-17)

1. Gender : Categorical

Transportation: Categorical

2. When

$$3. \quad a. \quad \frac{\text{girls who ride bus}}{\text{total \# of students}}$$

$$= \left( \frac{34}{188} \right) (100\%) = 18\%$$

$$b. \quad \frac{\text{girls who ride bus}}{\text{total \# of girls}}$$

$$= \left( \frac{34}{102} \right) (100\%) = 33\%$$

$$c. \quad \frac{\text{girls who ride bus}}{\text{total \# bus riders}}$$

$$= \left( \frac{34}{64} \right) (100\%) = 53\%$$

4. Look at totals of the variable they give you.

Male: 86 Females: 102

5. Find conditional frequencies (the percentages) for each transportation mode for males.

Males

$$\text{Bus } (30/86)(100\%) = 35\%$$

$$\text{Ride } (37/86)(100\%) = 43\%$$

$$\text{Personal } (19/86)(100\%) = 22\%$$

The most males ride to school, then bus, the personal.

(make a better sentence than this one.)

## Associations :

- variables can be associated with with each other in different ways to different degrees
- the best way to tell whether they are associated is to ask whether they are not

## Independence:

- When the distribution for one variable is almost the same for another variable, we say the variables are independent

6. Yes, transportation and gender are independent. The values for each gender/transportation combination are close enough to be independent.

## Worksheet Review: (p. 3-15)

1. Party: Rep., Dem., Other  $\rightarrow C$ Age: 18-29, 30-49, 50-64,  
65+  $\rightarrow C$ 

2. Who: Respondents

What: Age, Party

When: N/A

Where: N/A

Why: Understand party preference

How: Survey

$$3. \frac{\text{Republican}}{\text{Total respondents}} = \left( \frac{16535}{54384} \right) (100\%)$$
$$= 30\%$$

$$4. \frac{(\text{Under 30}) + (\text{over 65})}{\text{total respondents}}$$

$$= \left( \frac{10139 + 9784}{54384} \right) (100\%)$$

$$= 37\%$$

$$5. \frac{\text{under 30}}{\# \text{ of others}} = \left( \frac{4765}{20666} \right) (100\%)$$
$$= 23\%$$

$$6. \frac{\# \text{ of others under 30}}{\text{total } \# \text{ under 30}}$$
$$= \left( \frac{4765}{10139} \right) (100\%)$$
$$= 47\%$$

7. 18-29: 10139  
30-49: 21473  
50-64: 12988  
65+: 9784

8. Democrat

18-29  $(2738/17183)(100\%) = 16\%$   
30-49  $(6442/17183)(100\%) = 37\%$   
50-64  $(4286/17183)(100\%) = 25\%$   
65+  $(3718/17183)(100\%) = 22\%$

Because difference between largest percentage is 21% (37% - 16%), the variables are NOT independent.

9. Using the Democrats as an example, we see the variables are NOT independent.